

## 本周周报（2012.10.29-2012.11.4）

马晓红

本周工作：

### 1. 阅读了两篇论文：

**DBXplorer: A System for Keyword-Based Search over Relational Databases:**这篇文章也是一个关于在关系数据库上关键字检索系统。整体来说是先建立一个 **Symbol Table**，**Symbol Table** 用来记录每个关键字的位置，对于要查询的关键字先找出这些关键字所在的数据库中的相应的表和列或者具体的一个 **cell**，然后根据这些表之间的外键关系将这些表连起来构造 **SQL** 语句，检索出包含所有关键字的记录。上次也阅读了一篇在实体数据库上进行关键字检索的论文，那篇论文构造 **SQL** 语句的方法是通过构造查分查询对，在搜索引擎上查分查询对结果之间的差异体现出某个关键字对应的属性。

**Exploiting web search to generate synonyms for entities:** 这篇论文讲的是通过网络搜索来产生实体的同义词。采用的方法是构建表示该实体的关键字子集，在搜索引擎上搜索这些子集，在产生的文档中判断该关键字子集是否是该实体的同义词。这篇文章侧重于提高这种方法的检索效率，也就是探究如何减少利用搜索引擎搜索的次数。

### 2. 本周大部分时间用来完成数据挖掘的课程论文作业。

下周工作：

1. 王桂珍师姐已经回来了，看看下周找个时间交接一些文本分析的一些任务。
2. 应该要花一些时间来完成操作系统等课程的课程报告。
3. 继续看关于信息检索的论文。